Diffusion effects in gradient echo memory

X.-W. Luo,[1,2] J. J. Hope,[3] B. Hillman,[3] and T. M. Stace[1,*]

[1]*ARC Centre for Engineered Quantum Systems, University of Queensland, St Lucia, Queensland 4072, Australia*

[2]*Key Laboratory of Quantum Information, CAS, University of Science and Technology of China, Hefei, Anhui, 230026, People'sRepublic of China*

[3]*Department of Quantum Science, Research School of Physics and Engineering, Australian National University, Canberra, Australian Capital Territory 0200, Australia*

(Received 18 March 2013; published 24 June 2013)

We study the effects of diffusion on a Λ-gradient echo memory, which is a coherent optical quantum memory, using thermal gases. The efficiency of this memory is high for short storage time, but decreases exponentially due to decoherence as the storage time is increased. We study the effects of both longitudinal and transverse diffusion in this memory system, and give both analytical and numerical results that are in good agreement. Our results show that diffusion has a significant effect on the efficiency. Further, we suggest ways to reduce these effects to improve storage efficiency. We also report on a mechanism by which the rate of expansion of the transverse width of the beam is reduced compared to the naive expectation of diffusive effects, as observed in recent experiments.

DOI: 10.1103/PhysRevA.87.062328 PACS number(s): 03.67.Ac, 32.80.−t, 42.50.Md

I. INTRODUCTION

Quantum memory is an important tool in many quantum information protocols, including quantum repeaters for long-distance quantum communication [1] and identity quantum gates in quantum computation [2]. The required optical quantum memory must first be able to receive and recreate quantum light fields and second store multiple quantum states of light for on-demand recall with fidelity exceeding the classical limit.

Numerous optical quantum memories have been developed based on ensembles of material particles, including electromagnetically induced transparency (EIT)–based quantum memory [3,4], far-detuned Raman process memory [5], and photon-echo quantum memories: controlled reversible inhomogeneous broadening (CRIB) memory [6,7], atomic frequency combs (AFC) memory [8], and gradient echo memory (GEM) [9–11]. A review of these schemes can be found in Ref. [12]. EIT has achieved efficiencies of 43% [13] and has been used to store light pulses for multiple seconds [14]. AFC memory has achieved efficiencies of 35% [15] and have been used to store entangled states [16,17]. Besides the ensemble approach, single-atom quantum memories have also been developed [18,19]. Unlike the ensemble memories, where the information is stored in collective excitations, single-atom memories store single-photon states in the internal states of a single atom [18] or an atom-mirror system [19]. Efficiency attained by single-atom memory is over 9% [18], and teleportation between remote single-atom quantum memories has also been studied [20].

The highest efficiency attained for an experimental quantum memory is 87% by a Λ-GEM scheme [21] using warm rubidium vapor. The spatial multimode properties of a Λ-GEM have been examined experimentally [22]. In this paper, we examine the effects of atomic diffusion on the Λ-GEM system, which may limit this efficiency for larger storage times.

Λ-GEM is a memory using a three-level, Λ-type atom (Fig. 1). The input optical pulse couples the two metastable lower states through a control field. The excited state is coupled in the far detuned region, so the three-level atoms can be treated as effective two-level atoms. These effective two-level atoms have linearly increasing atomic Zeeman shifts along the length of the storage medium. The pulse is first absorbed, and then by simply reversing the sign of the magnetic field, the pulse is retrieved in the forward direction. The incident signal field is converted to a collective atomic excitation known as a spin wave, which is distributed as a function of position. The Brownian motion of the gaseous atoms will cause diffusion, which will disturb spatial coherence of the atomic spin wave, leading to decoherence. For a plane wave, only axial diffusion is important, but transverse diffusion becomes significant when a realistic beam profile is included. There has been recent interest in the effects of diffusion in the EIT [23–25].

In this work, we study the effects of diffusion in Λ-GEM system, giving analytical and numerical results. We suggest ways to reduce these effects to improve storage efficiency and also report on a mechanism that explains the discrepancy between the predicted diffusion coefficients based on buffer gas pressure calculation and the echo signal shape, which was seen in Ref. [22]. The effects of diffusion strongly depend on the spatial frequency of the spin wave. In the longitudinal direction, we find that the spatial frequency of the spin wave starts initially at a nonzero value, and then it will increase or decrease, depending on the sign of the gradient η. It is possible to arrange for the mean spatial frequency to approach zero after the absorption of the optical pulse, allowing the diffusion effects to be minimized by turning off the gradient. The control field with a realistic transverse Gaussian profile leads to a transverse variation in the phase of the spin wave, which under the influence of diffusion will reduce the apparent width of the emitted read-out signal.

II. Λ GRADIENT ECHO MEMORY

We consider a medium consisting of Λ-type three-level atoms with two metastable lower states as shown in Fig. 1.

*stace@physics.uq.edu.au

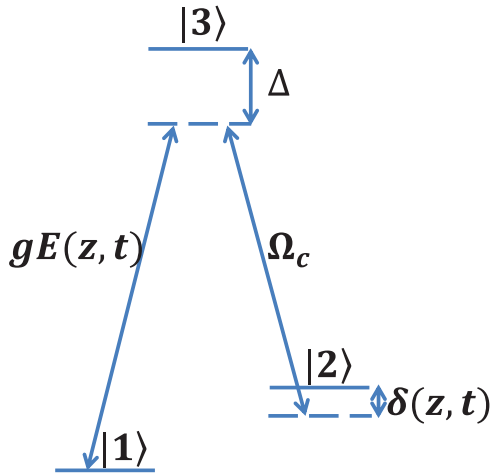

FIG. 1. (Color online) Level structure of Λ-type three-level atom.

The ground state $|1\rangle$ and the excited state $|3\rangle$ are coupled by a weak optical field, and the positive frequency component of the electric field is described by the slowly varying operator

$$\hat{E}(\mathbf{r},t)=\sum_{\mathbf{k}}\sqrt{\frac{1}{V}}a_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{r}}e^{-ik_0z}e^{i\omega_0t} \tag{1}$$

with detuning Δ, where V is the quantization volume, ω_0 is the carrier frequency of the quantum field, and $k_0=\omega_0/c$. The excited state $|3\rangle$ is also coupled to the metastable state $|2\rangle$ via a coherent control field with Rabi frequency Ω_c and a two-photon detuning δ. This two-photon detuning is spatially varied $\delta(z,t)=\eta(t)z$, with time-dependent gradient $\eta(t)$. Then the interaction Hamiltonian in the rotating frame with respect to the field frequencies is

$$\hat{H}=\sum_n\left\{\hbar\Delta\sigma_{33}^{(n)}+\hbar\delta(z_n,t)\sigma_{22}^{(n)}+\hbar\sum_{\mathbf{k}}\left[\hbar g_{\mathbf{k}}a_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{r}_n}\sigma_{31}^{(n)}+\hbar\Omega_c(\mathbf{r}_n)\sigma_{32}^{(n)}+\text{H.c.}\right]\right\}, \tag{2}$$

where $g_{\mathbf{k}}=\wp\sqrt{\frac{\omega_{\mathbf{k}}}{2\hbar\epsilon_0V}}$ is the atom-field coupling constant with $\wp$ being the dipole moment of the 1–3 transition and $\sigma_{\mu\nu}^{(n)}=|\mu\rangle_n\langle\nu|$ is an operator acting on the nth atom at $\mathbf{r}_n=(x_n,y_n,z_n)$. We assume that initially all atoms are in their ground state $|1\rangle$. We transform to collective operators, which are averages over atomic operators over a small volume centered at $\mathbf{r}$ containing $N_{\mathbf{r}}\gg1$ particles,

$$\sigma_{\mu\nu}(\mathbf{r},t)=\frac{1}{N_{\mathbf{r}}}\sum_{j=1}^{N_{\mathbf{r}}}\sigma_{\mu\nu}^{(j)}(t). \tag{3}$$

From the Heisenberg-Langevin equations in the weak probe region ($\sigma_{11}\simeq1,\sigma_{22}\simeq\sigma_{33}\simeq0$), we get the Maxwell-Bloch equations [26],

$$\begin{aligned}
\dot{\sigma}_{13}^{(n)}&=-(\gamma_{13}+i\Delta)\sigma_{13}^{(n)}+ige^{ik_0z_n}E(\mathbf{r}_n,t)+i\Omega_ce^{ik_cz_n}\sigma_{12}^{(n)},\\
\dot{\sigma}_{12}^{(n)}&=-[\gamma_{12}+i\delta(z_n,t)]\sigma_{12}^{(n)}+i\Omega_ce^{-ik_cz_n}\sigma_{13}^{(n)},\\
\left(\frac{\partial}{\partial t}+c\frac{\partial}{\partial z}-ic\frac{\nabla_x^2+\nabla_y^2}{2k_0}\right)&E(\mathbf{r},t)=igNe^{-ik_0z}\sigma_{13}(\mathbf{r},t),
\end{aligned} \tag{4}$$

where $\gamma_{\nu\mu}$ are the decay rates, $g=\wp\sqrt{\frac{\omega_0}{2\hbar\epsilon_0}}$, and N is the atomic density. We have assumed that $g_{\mathbf{k}}\simeq\wp\sqrt{\frac{\omega_0}{2\hbar\epsilon_0V}}$ and $\Omega_c(\mathbf{r})=\Omega_ce^{ik_cz}$. We also omit the Langevin noise operators since here we are more interested in the decoherence caused by diffusion. This is equivalent to making a semiclassical approximation for the electric field and the atomic coherences.

In Eq. (4), $ic\frac{\nabla_x^2+\nabla_y^2}{2k_0}$ is the diffraction term, and generally, the diffraction effects can be neglected [27]. Notice that we are here considering the regime $t_p\gg L/c$, where $2t_p$ is the temporal width of the signal, and $2L$ is the length of the medium. This allows us to neglect temporal retardation effects; i.e., we can neglect the temporal derivative in the third equation of Eq. (4). Also, since the atoms are far detuned ($\Delta\gg\gamma_{13},\Omega_c$), we adiabatically eliminate the fast oscillations and set $\dot{\sigma}_{13}^{(n)}=0$. Then we have $\sigma_{13}=(ge^{ik_0z}E+\Omega_ce^{ik_cz}\sigma_{12})/\Delta$, and we get the reduced Maxwell-Bloch equations,

$$\begin{aligned}
\dot{\sigma}_{12}^{(n)}&=-\left[i\delta(z_n,t)-i\frac{\Omega_c^2}{\Delta}\right]\sigma_{12}^{(n)}\\
&\quad+i\frac{g\Omega_c}{\Delta}e^{i(k_0-k_c)z_n}E(\mathbf{r}_n,t),\\
\frac{\partial}{\partial z}E(\mathbf{r},t)&=i\frac{gN\Omega_c}{c\Delta}e^{-i(k_0-k_c)z}\sigma_{12}(\mathbf{r},t)+i\frac{g^2N}{c\Delta}E(\mathbf{r},t).
\end{aligned} \tag{5}$$

Here we neglect decay, i.e., $\gamma_{12}\to0$, since we consider the storage time much less than $1/\gamma_{12}$.

III. DIFFUSION

We now consider the effects of diffusion on the atomic state. In order to isolate the motional effects of diffusion from collisional dephasing, we assume that the collisions between atoms do not change the state of the atom. Then we derive the diffusion equation for the atomic density matrix ρ. Space is divided into volume elements with length Δr and center r. We associate a density matrix $\rho(r,t)$ with atoms in this volume element, given by

$$\rho(r,t)=\frac{1}{N_r}\sum_{j=1}^{N_r}\rho^{(j)}(t),$$

where N_r is the atom number in volume centered at r. The total density matrix for the entire system is assumed to be the tensor product of these local density matrices.

Diffusion causes an exchange of atoms between adjacent volumes. During a short time Δt, a fraction ϵ of the atoms in slice r migrate into slice $r\pm\Delta r$. There is also atomic flux back into slice r from $r\pm\Delta r$. We assume that the total number density of the atoms is uniform, so the state at r and $t+\Delta t$ is described by the new density matrix which is the average of the density matrix of atoms remaining in the volume and those that have migrated in to it. The diffusive component of the evolution is therefore

$$\begin{aligned}
&\rho(r,t+\Delta t)\\
&\quad=(1-2\epsilon)\rho(r,t)+\epsilon[\rho(r+\Delta r,t)+\rho(r-\Delta r,t)]\\
&\quad\Rightarrow\partial_t\rho(r,t)=D\nabla^2\rho(r,t),
\end{aligned} \tag{6}$$

where $D = \epsilon \Delta r^2/\Delta t$ is the diffusion coefficient. With the same consideration, we get the diffusive component evolution for the atomic correlation functions:

$$\dot{\sigma}_{\mu\nu}(r,t) = D\nabla^2\sigma_{\mu\nu}(r,t). \tag{7}$$

Now we introduce the interaction with optical fields. Since diffusion is caused by Brownian motion, this will lead to Doppler shifts in the various detunings. We consider the interaction between the optical field and a single atom, and quantify the effects of these Doppler shifts. The atom moves at some random velocity, and there will be a Doppler shift for both the signal and control fields, so the detunings in Eq. (4) become $\Delta = \Delta_0 + \Delta_{\rm Dopp}$, and $\delta = \delta_0 + \delta_{\rm Dopp}$, with Δ_0, δ_0 the detunings for stationary atoms and $\Delta_{\rm Dopp}, \delta_{\rm Dopp}$ the Doppler shifts. Typically, the one-photon Doppler shift $\Delta_{\rm Dopp} \ll \Delta_0$, and state $|3\rangle$ is still far detuned. The adiabatic elimination is still valid in the presence of the Brownian motion–induced Doppler shift, and we can still reduce the three-level atom to an effective two-level atom. The Maxwell-Bloch equation will still reduce to Eq. (5), but with one-photon detuning $\Delta = \Delta_0 + \Delta_{\rm Dopp}$ and two-photon detuning $\delta = \delta_0 + \delta_{\rm Dopp}$. So, for the reduced two-level atomic system, the diffusive Maxwell-Bloch equation for the collective correlation $\sigma_{12}(z,t)$ averaged over atoms in each volume is

$$\begin{aligned}\dot{\sigma}_{12}(\mathbf{r},t) &= i\frac{g\Omega_c}{\Delta}e^{i(k_0-k_c)z}E(\mathbf{r},t)\\ &\quad - i\delta(z,t)\sigma_{12}(\mathbf{r},t) + D\nabla^2\sigma_{12}(\mathbf{r},t),\\ \frac{\partial}{\partial z}E(\mathbf{r},t) &= i\frac{gN\Omega_c}{c\Delta}e^{-i(k_0-k_c)z}\sigma_{12}(\mathbf{r},t) + i\frac{g^2N}{c\Delta}E(\mathbf{r},t).\end{aligned} \tag{8}$$

We have absorbed the Stark shift $\frac{\Omega_c^2}{\Delta}$ into the two-photon detuning. Here our diffusive Maxwell-Bloch equation is consistent with the result in the EIT system [23,25].

Notice that the signal and control fields are copropagating, so the Doppler broadening width for δ is typically 1 kHz, which is much smaller than the frequency width of the signal field ($\sim$1 MHz), so we neglect this two-photon Doppler broadening $\delta_{\rm Dopp}$ and replace $\delta(z,t)$ by $\delta_0(z,t) = \eta(t)z, z \in [-L,L]$.

For the one-photon detuning $\Delta = \Delta_0 + \Delta_{\rm Dopp}$, after we make the adiabatic elimination, it will appear in the denominator [see Eq. (8)], so

$$\frac{1}{\Delta} \simeq \frac{1}{\Delta_0}\left[1 - \frac{\Delta_{\rm Dopp}}{\Delta_0} + \left(\frac{\Delta_{\rm Dopp}}{\Delta_0}\right)^2\right].$$

The linear term in $\Delta_{\rm Dopp}$ will vanish when we average over many atoms in a volume centered at r, so we can replace Δ by Δ_0 in our Maxwell-Bloch equation, with second-order accuracy [typically $(\Delta_{\rm Dopp}/\Delta_0)^2 \sim 10^{-3}$].

IV. ANALYTIC CALCULATION AND NUMERICAL SIMULATION

To quantify the effects of diffusion, we solve for the atomic dynamics. There are three distinct phases during the storage: write-in $-t_0 < t < 0$, during which the signal is absorbed by the memory; hold $0 < t < t_H$, during which the information is stored in the memory and the gradient is turned off; and read-out $t_H < t < t_H + t_0$, during which the signal is emitted by turning on the flipped gradient.

We quantify the effects of diffusion by the read-out efficiency ε defined to be

$$\varepsilon = \frac{\int_{t_H}^{t_H+t_0}|f_{\rm out}(t)|^2dt}{\int_{-t_0}^{0}|f_{\rm in}(t)|^2dt}, \tag{9}$$

where $f_{\rm out}(t) = E(z = L, t > t_H)$ is the output field and $f_{\rm in}(t) = E(z = -L, t < 0)$ is the input field. We solve for $f_{\rm out}(t)$ both numerically and analytically, and consider the effects of diffusion in axial (longitude) and radial (transverse) directions separately.

A. Longitudinal diffusion

For a uniform plane wave, transverse diffusion is irrelevant. We replace $\mathbf{r}$ by z in Eq. (8) and consider the longitude diffusion in a one-dimensional model. Now the Maxwell-Bloch equation is

$$\begin{aligned}\dot{\sigma}_{12}(z,t) &= i\frac{g\Omega_c}{\Delta}e^{i(k_0-k_c)z}E(z,t)\\ &\quad - i\delta(z,t)\sigma_{12}(z,t) + D\nabla_z^2\sigma_{12}(z,t),\\ \frac{\partial}{\partial z}E(z,t) &= i\frac{gN\Omega_c}{c\Delta}e^{-i(k_0-k_c)z}\sigma_{12}(z,t) + i\frac{g^2N}{c\Delta}E(z,t).\end{aligned} \tag{10}$$

We now investigate the longitude diffusion effects during the write-in process, the hold time, and the read-out processes separately.

To compute $f_{\rm out}(t)$, we evolve Eq. (10) using η [as in Fig. 5(a)]. Following the method given in Ref. [9], we first propagate $E(z,t)$ and $\sigma_{12}(z,t)$ forward with boundary condition $E(z = -L, t < 0) = f_{\rm in}(t)$ to find their values at time t_H. Then we propagate E and σ_{12} backward to time t_H, with final condition $E(z = L, t > t_H) = f_{\rm out}(t)$, and solve for $f_{\rm out}(t)$ by matching the two solutions at time t_H.

1. Write

By considering the diffusion effects during the write-in process, we find that (see Appendix A)

$$f_{\rm out}(t_H + t) = e^{\frac{-D}{3\eta}[k_i^3 - (k_i - \eta t)^3]}f_{\rm in}(-t)\bar{G}, \tag{11}$$

where $k_i = \frac{g^2N}{c\Delta} + k_0 - k_c - \frac{\beta}{L}$ is the initial spatial frequency of $\sigma_{12}(z,t)$, and

$$\bar{G} = \left|\eta L\left(t + \frac{\beta}{\eta L}\right)\right|^{-i2\beta} e^{i\frac{2Lg^2N}{c\Delta}} e^{-i\frac{g_{\rm eff}^2N}{c\eta(t+\frac{\beta}{\eta L})}t_H}\Gamma(i\beta)/\Gamma(-i\beta)$$

is a phase factor, with $g_{\rm eff} = g\Omega_c/\Delta$ and $\beta = \frac{g_{\rm eff}^2N}{\eta c}$. For a pulse with Gaussian temporal profile $f_{\rm in} = Ae^{-(t+t_{\rm in})^2/t_p^2}$, we find

$$\varepsilon_W = \frac{\int_{-t_0}^{0}dte^{-2(t+t_{\rm in})^2/t_p^2}e^{\frac{-2D}{3\eta}[k_i^3-(k_i+\eta t)^3]}}{\int_{-t_0}^{0}dte^{-2(t+t_{\rm in})^2/t_p^2}}. \tag{12}$$

Typically, $D\eta^2t_p^3$ is very small, and then the efficiency is

$$\varepsilon_W = \sqrt{\alpha_W}e^{-\tau_W} + O\left(D^2\eta^4t_p^6\right), \tag{13}$$

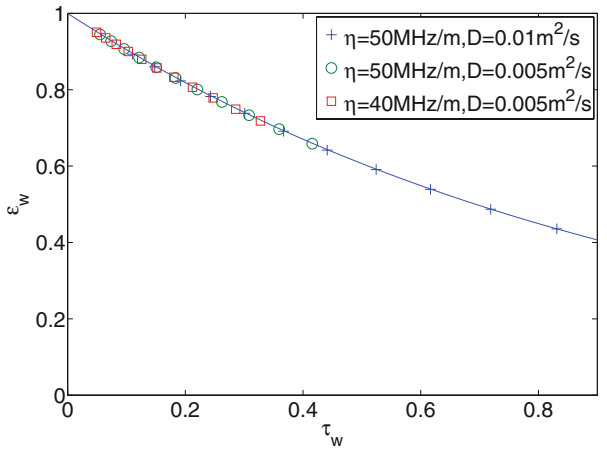

FIG. 2. (Color online) The efficiency decay with respect to the dimensionless parameter τ_W for longitude diffusion during the write-in time; the points are numerical results, and the curve is Eq. (14).

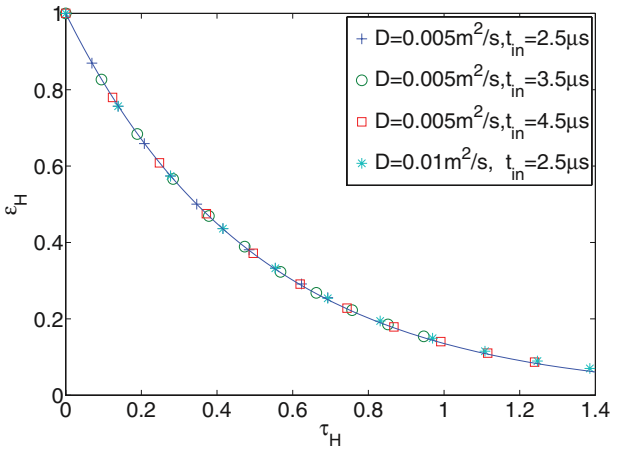

FIG. 3. (Color online) The efficiency decay with respect to the dimensionless parameter τ_H for longitude diffusion during hold time; the points are numerical results, and the curve is Eq. (17).

where $\alpha_W = \frac{1}{1-D\eta^2 t_p^2(k_i/\eta - t_{\rm in})}$ and $\tau_W = \frac{2D\eta^2}{3}[(\frac{k_i}{\eta})^3 - (\frac{k_i}{\eta} - t_{\rm in})^3]$ are dimensionless parameters. For typical experimental parameters, $\alpha_W \simeq 1$, then

$$\varepsilon_W \simeq e^{-\tau_W}. \tag{14}$$

We also numerically solve Eq. (10) with diffusion during the write-in process, using XMDS [28]. XMDS is a cross-platform, open source package for numerically integrating initial value problems that range from a single ordinary differential equation up to systems of coupled stochastic partial differential equations. We calculate the efficiency for different values of the diffusion rate D, input time $t_{\rm in}$, etc. The results are shown in Fig. 2 [points are numerical results, and the curve is Eq. (14)]. We plot the efficiency ε_W with respect to the rescaled dimensionless parameter τ_W, so all the points with different parameters collapse on a single curve.

2. Hold

During the storage time $[0,t_H]$, we find (see Appendix A)

$$f_{\rm out}(t_H + t) = e^{-Dt_H(k_i-\eta t)^2} f_{\rm in}(-t)\bar{G}. \tag{15}$$

For the above Gaussian shape input, the efficiency is given by

$$\varepsilon_H = \sqrt{\alpha_H} e^{-2\alpha_H \tau_H}, \tag{16}$$

where $\alpha_H = \frac{1}{t_p^2}/(\frac{1}{t_p^2} + Dt_H\eta^2)$ and $\tau_H = Dt_H k_H^2$ are dimensionless parameters, with $k_H = k_i - \eta t_{\rm in}$. For typical experimental parameters, $\alpha_H \simeq 1$ and we have

$$\varepsilon_H \simeq e^{-2\tau_H}. \tag{17}$$

We also numerically solve Eq. (10) with diffusion during hold time, using XMDS. We calculate the efficiency for different values of the diffusion rate D, storage time t_H, etc. The results are shown in Fig. 3 [points are numerical results, and the curve is Eq. (17)]. We plot the efficiency ε_H with respect to the rescaled dimensionless parameter τ_H, so all the points with different parameters collapse on a single curve.

3. Read

The diffusion effects during the read-out process are the same as the diffusion effects of the write-in process (see Appendix A), so we simply have

$$\varepsilon_R = \varepsilon_W.$$

B. Transverse diffusion

We now quantify the effects of diffusion for a beam with realistic transverse Gaussian profile. The efficiency for a three-dimensional model is defined as

$$\varepsilon = \frac{\int |f_{\rm out}(x,y,t_H+t)|^2 dxdydt}{\int |f_{\rm in}(x,y,t)|^2 dxdydt}. \tag{18}$$

Equation (8) can be solved in Fourier space k_x,k_y, and notice that

$$\varepsilon = \frac{\int |f_{\rm out}(k_x,k_y,t_H+t)|^2 dk_x dk_y dt}{\int |f_{\rm in}(k_x,k_y,t)|^2 dk_x dk_y dt}. \tag{19}$$

Equation (8) can be reduced to a quasi-one-dimensional (quasi-1D) problem and can be solved as before (see Appendix B). For transverse diffusion, the output pulse will be

$$f_{\rm out}(k_x,k_y,t_H+t) = e^{-2\gamma_k t} e^{-\gamma_k t_H} f_{\rm in}(k_x,k_y,-t)\bar{G}, \tag{20}$$

where $\gamma_k = D(k_x^2 + k_y^2)$.

If the input pulse has both Gaussian temporal and transverse profile,

$$f_{\rm in}(x,y,t) = Ae^{-(x^2+y^2)/a^2} e^{-(t+t_{\rm in})^2/t_p^2},$$

then $\gamma_k t_p \sim Dt_p/a^2$, which is typically small. Thus the memory efficiency is

$$\varepsilon_\perp = \frac{1}{1+\tau_\perp} + O\left(\gamma_k^2 t_p^2\right), \tag{21}$$

where $\tau_\perp = 4D(t_H + 2t_{\rm in})/a^2$ is a dimensionless parameter.

We also numerically solve Eq. (8) with $\nabla^2 = \nabla_x^2 + \nabla_y^2$. We calculate the efficiency for different values of a, t_H, etc. The results are shown in Fig. 4 [points are numerical

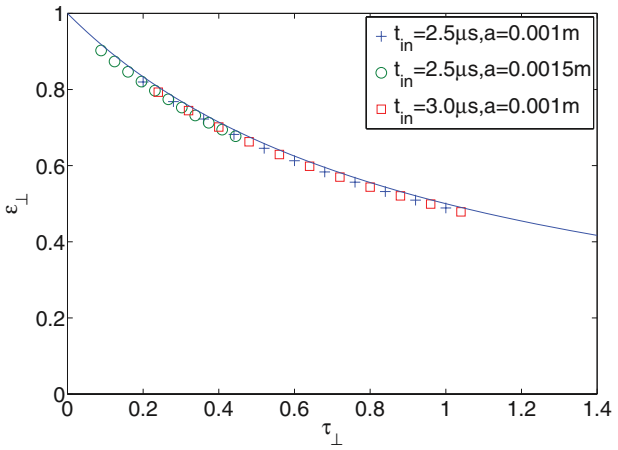

FIG. 4. (Color online) Efficiency decay with respect to $\tau_\perp$ for transverse diffusion; the points are numerical results, and the curve is Eq. (21).

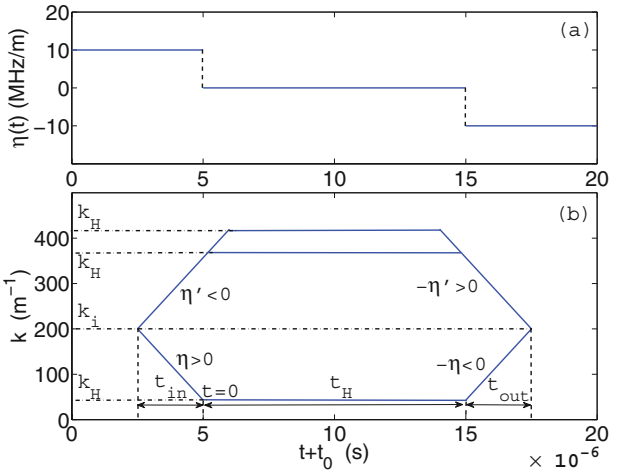

FIG. 5. (Color online) (a) The gradient is turned off during hold time and flipped for read-out. (b) An illustration of the spatial frequency for $\sigma_{12}(z,t)$, k decreases or increases depending on the sign of the gradient.

results, and the curve is Eq. (21)]. We plot the efficiency $\varepsilon_\perp$ with respect to the rescaled dimensionless parameter $\tau_\perp$, so all the points with different parameters collapse on a single curve.

C. Total diffusion

Experimentally, longitude and transverse diffusion coexist during the whole process. Combining all the diffusive contributions mentioned above, we get the output field as (Appendix B)

$$f_{\text{out}}(k_x,k_y,t_H+t) = e^{\frac{-2D}{3\eta}[k_i^3-(k_i-\eta t)^3]}e^{-Dt_H(k_i-\eta t)^2} \times e^{-2\gamma_k t-\gamma_k t_H} f_{\text{in}}(k_x,k_y,-t)\bar{G}. \quad (22)$$

We consider input pulse with both Gaussian temporal and transverse profile as above; typically, $D\eta^2 t_p^3, \gamma_k t_p$ are very small. Then the total efficiency will be

$$\varepsilon_{\text{tot}} = \sqrt{\frac{1}{1/\alpha_H+2/\alpha_W-2}}e^{-2\tau_W}e^{-2\tau_H}\frac{1}{1+\tau_\perp} + O\left[\left(D\eta^2 t_p^3,\gamma_k t_p\right)^2\right]. \quad (23)$$

Typically, $\alpha_H \simeq 1, \alpha_W \simeq 1$, so we have

$$\varepsilon_{\text{tot}} \simeq \varepsilon_W \varepsilon_H \varepsilon_R \varepsilon_\perp.$$

D. Efficiency optimization and estimation

Our model did not examine other decoherence processes, such as control-field-induced scattering and ground-state decoherence. Our results simply quantify the effects of motional diffusion on GEM efficiency and therefore the represent upper estimates for the performance of GEM.

Structures with larger spatial frequency will decay faster under diffusion. For the 1D model during hold time, we have (see Appendix A)

$$\sigma_{12}(k,t) \propto f_{\text{in}}\left(\frac{k-k_i}{\eta}\right). \quad (24)$$

The input pulse is centered at $-t_{\text{in}}$, and then we have $k \sim k_H = k_i - \eta t_{\text{in}}$. We see that k will increase or decrease as t_{in} increases, depending on the sign of η. During the hold time $[0,t_H]$, the gradient is turned off, and k will hold its value. The read-out process is symmetric to the write-in process (see Fig. 5), returning the quasimomentum to its original distribution. Figure 6 shows the numerical result by solving Maxwell-Bloch equations.

Notice that the initial k vector starts at a nonzero value k_i, which is mainly determined by the wave vector difference between the probe and control field in the medium. k_i could be positive, negative, or zero, depending on the relative momentum of the control and probe fields in the medium. Since structures with larger $|k|$ will decay faster under diffusion, we aim to have very small $|k|$ after the writing process. The k vector will evolve linearly during the writing process due to the gradient, so the storage time will be improved by starting with a positive (negative) writing gradient η for a positive (negative) k_i. As noted below, $|k_i| > \frac{1}{L}$ is required to get $k_H = 0$. Keeping this spatial frequency small also requires the removal of the

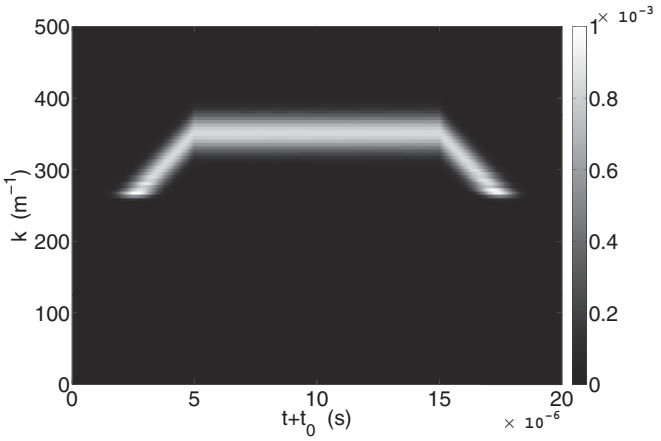

FIG. 6. Numerical results of $|\sigma_{12}(k,t)|$ with write-in gradient $\eta' < 0$.

gradient during the storage part of the process and only turning on the flipped gradient during read-out.

For realistic $t_{\rm in}$ and η, we can get zero spatial frequency $k_H = 0$, for which $\tau_H = 0$, and this will minimize the diffusional decay rate. We then have $\varepsilon_H = \sqrt{\alpha}$ and $\varepsilon_W = e^{-2Dk_i^2 t_{\rm in}/3}$. Including transverse diffusion, we get the total efficiency for input field with transverse Gaussian profile

$$\varepsilon_{\rm tot} = e^{-4Dk_i^2 t_{\rm in}/3}\sqrt{\alpha}\frac{1}{1+4D(t_H+2t_{\rm in})/a^2}. \tag{25}$$

The efficiency can be improved further by choosing a larger transverse width a; i.e, the effects of transverse diffusion will be reduced by using a smooth field in the transverse direction.

We note that the circumstances in which a GEM will be useful are those for which all dephasing, including that due to diffusion, is small. In this limit, a useful approximate expression for the GEM efficiency is given by

$$\varepsilon_{\rm tot} \simeq 1 - \frac{4Dk_i^2 t_{\rm in}}{3} - \frac{Dt_H\eta^2 t_p^2}{2} - \frac{4D(t_H+2t_{\rm in})}{a^2}, \tag{26}$$

as the inefficiencies arising from each diffusive process considered above add together.

Experimental considerations give estimates of the achievable GEM efficiency. In particular, to ensure the bandwidth of the memory is large enough to absorb the input field, we require $|\eta t_p| > \frac{1}{L}$, and $t_{\rm in} > t_p$ to ensure that the whole pulse enters the medium during the write-in process; also $|k_i| > \frac{1}{L}$ is required to satisfy $k_H = 0$. So

$$\begin{aligned}\varepsilon_{\rm tot} &\lesssim 1 - \frac{4Dk_i^2 t_p}{3} - \frac{Dt_H}{2L^2} - \frac{4D(t_H+2t_p)}{a^2} \\ &\lesssim 1 - \frac{4Dt_p}{3L^2} - \frac{Dt_H}{2L^2} - \frac{4D(t_H+2t_p)}{a^2}.\end{aligned} \tag{27}$$

This gives a reasonable upper bound on the GEM efficiency, given the pulse duration $2t_p$, the hold time t_H, and the vapor length L, and beam width a.

1. Experimental considerations

Typical system parameters are $\omega_0 = 2\pi \times 377.107\,46$ THz, $\omega_0 - \omega_c = 2\pi \times 6.8$ GHz, $\Delta = -2\pi \times 1.5$ GHz, $\Omega_c \simeq 2\pi \times 20$ MHz, $g \simeq 2\pi \times 4.5$ Hz, $t_p = 1$ μs, $a \simeq 1.45$ mm, $2L = 0.2$ m, $\eta \simeq -2\pi \times 10$ MHz/m, and $N \simeq 0.5 \times 10^{18}$ m^{-3} [21,22,29]. The optical depth $|\beta| \simeq 3.8$ is sufficiently large. According to the formula in Ref. [30], we have $D \sim 0.004$ m^2/s for Rubidium atoms in buffer gas [21,22].

With these parameters, the diffusive decay is dominated by transverse diffusion. For example, for $t_{\rm in} = 5$ μs and $t_H = 0$, the maximum achievable efficiency is $\varepsilon_{\rm tot} \simeq 93\%$ ($\varepsilon_\perp \simeq \varepsilon_{\rm tot}$, $\varepsilon_H = 1$, $\varepsilon_W \simeq 1$).

We examine the input of the (1,1) Hermite-Gaussian mode $f_{\rm in}(x,y,t) \propto xy\, e^{-(x^2+y^2)/a^2}$ as an example of a higher order Hermite-Gaussian mode transverse profile. From Eqs. (18) and (20), we have $\varepsilon_\perp = (\frac{1}{1+\tau_\perp})^3$, and the longitude diffusion effects are the same as the Gaussian profile (i.e., the (0,0) Hermite-Gaussian mode). Thus, for diffusive decays, we have

$$\frac{\varepsilon_{(11)}}{\varepsilon_{(00)}} \propto \left(\frac{1}{1+\tau_\perp}\right)^2, \tag{28}$$

where $\varepsilon_{(ij)}$ is the read-out efficiency for (ij) Hermite-Gaussian mode. We find that the efficiency decays faster for higher order modes, and the ratio Eq. (28) decreases when the storage time increases. This is in agreement with experimental investigations [22].

E. Output beam width

After some storage time, transverse diffusion will tend to smear the spin wave density in the radial direction. Intuitively, we would expect this to lead to a spatially wider output beam than would be the case in the absence of diffusion.

This is certainly the case when the control field is radially uniform. To see this, we define the intensity distribution for the read-out signal as

$$I(r_\perp) = \int |f_{\rm out}(r_\perp, t_H + t)|^2 dt. \tag{29}$$

We suppose that the control field is turned off during the hold time, $[0,t_H]$, to avoid control field-induced scattering and that the gradient is always on and flipped at $t = 0.5t_H$. Also for typical experimental parameters, the effects of longitudinal diffusion is very weak, so we focus on transverse diffusion. We solve the Maxwell-Bloch equation using the same method as before. For a signal with a Gaussian transverse profile, we find that

$$I(r_\perp) \propto e^{r_\perp^2/[a^2+4D(2t_{\rm in}+t_H)]} \tag{30}$$

with $r_\perp^2 = x^2 + y^2$. Defining $w_{r_\perp}$ as the width of the output field, we have

$$w_{r_\perp}^2 = \frac{a^2}{4} + D(2t_{\rm in} + t_H), \tag{31}$$

which increases linearly with storage time (Fig. 9), at a rate determined by the diffusion coefficient.

Somewhat surprisingly, the experimentally measured rate of expansion of the read-out signal is smaller than that expected from atomic diffusion by a factor of 2 to 3 [22]. One possible explanation for this is the signal diffraction; as suggested in Ref. [22], diffusion leads to a beam with reduced divergence and the measurement is taken downstream. However in this experiment the scale of experimental setup is much smaller than the Rayleigh range, so the diffraction effect is too small to explain the observed discrepancy.

Instead, we find that the anomalously narrow output beam width can be explained by considering the control field with realistic transverse Gaussian profile. This leads to a transverse variation in the phase of the spin wave, which under the influence of diffusion leads to lower emission efficiency in the wings of the spin wave.

To analyze this effect quantitatively, we consider a Gaussian transverse variation in the control field, $\Omega_c(r_\perp) = \Omega e^{-r_\perp^2/w_c^2}$, with beam waist w_c. Then the two-photon detuning, δ, and the optical depth become $r_\perp$ dependent.

From our solution for the spin wave [see Appendix A, Eq. (A5)], we find that the inhomogeneity of the control field intensity will introduce a transverse variation in δ, which leads to a transverse dependence in the phase of the spin wave. Likewise, the transverse variation in the optical depth leads to a radially dependent longitudinal shift in the spin

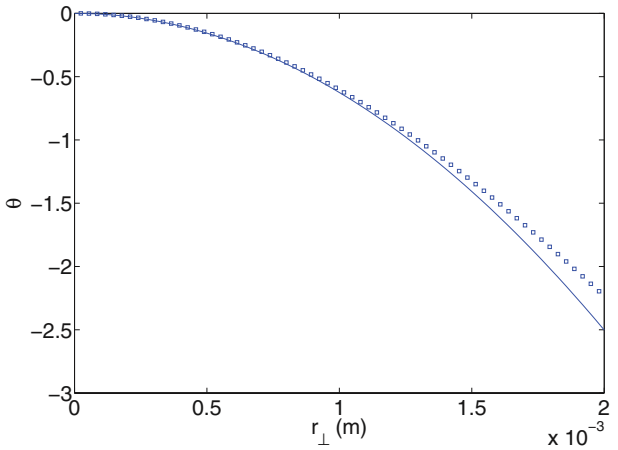

FIG. 7. (Color online) The extra phase θ for control field with Gaussian profile. Points are numerical results using typical parameters given in the main text, and the curve is the approximate expression in Eq. (32).

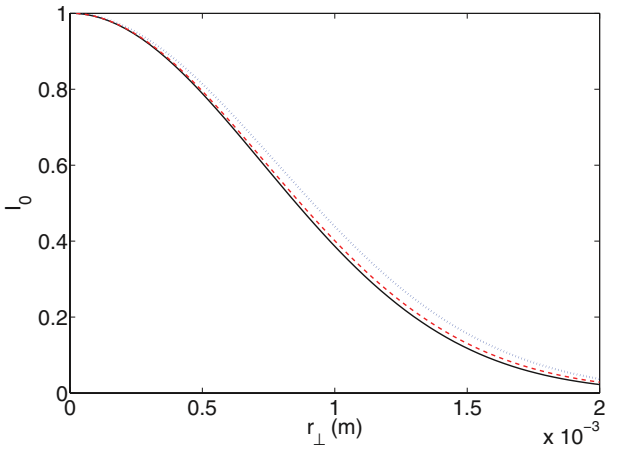

FIG. 8. (Color online) The intensity distribution for the read-out signal with $t_H = 16$ μs. To see the expansion clearly, we have renormalized the maximum of $I(r_\perp)$ to 1, and I_0 is the renormalized intensity distribution. The black solid curve is the input signal, the dotted one (blue online) is the read-out signal for the homogeneous control field, and the dashed one (red online) is the read-out signal for the control field with Gaussian profile.

wave $\sigma_{12}(\mathbf{r},t)$. In combination, these give rise to a radially dependent phase on the spin wave, $e^{i\theta(r_\perp)}$, with the effect of the control field typically being dominant. We compare solutions for this inhomogeneous control field with solutions for the homogeneous control field to obtain the phase difference $\theta(r_\perp)$ during hold time. Typically, the width of the control field, w_c, is much larger than the width of the signal field, a, so $\theta(r_\perp)$ is approximately quadratic in $r_\perp/w_c$:

$$\theta(r_\perp) = \left[-\frac{2\Omega^2 t_{\rm in}}{\Delta} + 2\beta \ln\left(\left| \frac{\eta L t_{\rm in}}{\beta} + 1 \right| \right) + 2\beta \left(1 - \frac{\Omega^2}{\Delta \eta L}\right) \left(\frac{\beta}{\eta L t_{\rm in} + \beta} + \frac{z}{L} \right) \right] \frac{r_\perp^2}{w_c^2}, \quad (32)$$

where β is the optical depth corresponding to Ω. Because of this quadratic phase variation across the spin wave, diffusion acts to wash out the spin-wave coherence more quickly at larger radius, so the read-out efficiency is suppressed at larger $r_\perp$. This tends to reduce the apparent width of the emitted read-out signal.

1. Experimental considerations

In the experimental results reported in Ref. [22], $w_c \simeq 3$ mm, and $t_{\rm in} \simeq 2$ μs. Using these parameters, Fig. 7 shows the transverse variation in the phase of the spin wave at $(z = 0, t = 0.5t_H)$ in the absence of diffusion. When diffusion is introduced, this transverse phase variation is smeared out, leading to reduced read-out efficiency in the wings of the spin-wave. Figure 8 compares the numerical results for the expansion of the read-out signal after a specific hold time, $t_H = 16$ μs, with a homogeneous control field (dotted, blue online) and with a spatially varying control field (dashed, red online), assuming the diffusion rate $D = 0.004$ m^2/s. Figure 9 shows the variation in the width of the output field as a function of hold time. We see that the expansion is slowed for a control field with Gaussian profile (squares), compared to the case of a uniform control field (circles). Importantly, this corresponds to a reduction of the beam width expansion rate by a factor of 2. The apparent diffusion rate extracted from this slower expansion rate is $D_{\rm eff} \simeq 0.002$ m^2/s. This is quantitatively in agreement with the observations in Ref. [22].

V. SUMMARY

We have studied the effects of diffusion on the efficiency of the Λ-gradient echo memory, both numerically and analytically. We find that the efficiency is dependent on the spatial frequencies k for both longitude diffusion and transverse diffusion: Higher k leads to more pronounced diffusive effects and reduced efficiency, as expected. We show that the storage efficiency can be improved by appropriate choice of the gradient during the hold phase.

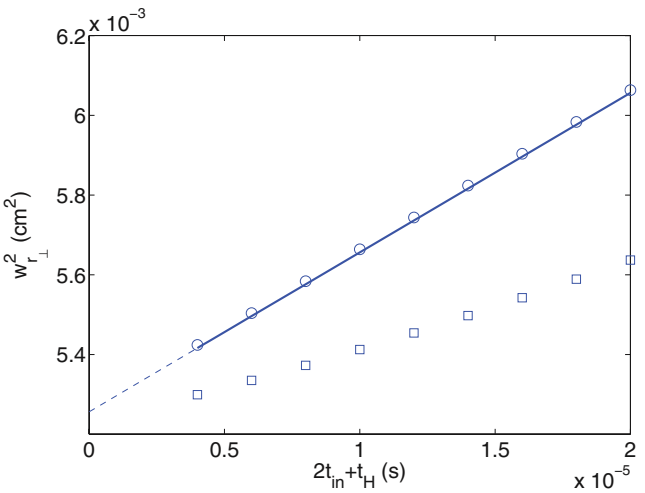

FIG. 9. (Color online) Expansion of the read-out signal. Circles are numerical results for the homogeneous control field. Squares are numerical results for the control field with a Gaussian profile, and the solid line is Eq. (31) for the homogeneous control field.

We established a mechanism by which the rate of expansion of the transverse width of the beam is reduced compared to the naive expectation of diffusive effects. This mechanism arises from the effects of diffusion on the transverse variation in the spin wave phase. We showed that with an experimentally reasonable choice of parameters, the magnitude of this effect is the same as that observed in recent experiments. When the density of the buffer gas in increased, the collision rate increases, leading to a smaller diffusion rate. However, this will lead to collision-induced dephasing, which will dominate at sufficiently high buffer gas pressures. This implies a trade-off between diffusion- and collision-induced dephasing. This will be the subject of future research.

ACKNOWLEDGMENTS

J. Hope and B. Hillman thank M. Hosseini, D. Higginbottom, O. Pinel, and B. Buchler for helpful discussions of the modeling and experiments. J. Hope was supported by the ARC Future Fellowship Scheme. X.-W. Luo thanks G. J. Milburn for helpful discussions and gratefully acknowledges the National Natural Science Foundation of China (Grant No. 11174270), the National Basic Research Program of China (Grant No. 2011CB921204), and CAS for financial support and the University of Queensland for kind hospitality.

APPENDIX A

The Maxwell-Bloch equation for the one-dimensional model is

$$\begin{aligned}\dot{\sigma}_{12}(z,t) &= i\frac{g\Omega_c}{\Delta}e^{i(k_0-k_c)z}E(z,t)\\ &\quad - i\delta(z,t)\sigma_{12}(z,t) + D\nabla_z^2\sigma_{12}(z,t), \qquad \text{(A1)}\\ \frac{\partial}{\partial z}E(z,t) &= i\frac{gN\Omega_c}{c\Delta}e^{-i(k_0-k_c)z}\sigma_{12}(z,t) + i\frac{g^2N}{c\Delta}E(z,t).\end{aligned}$$

To find the solution during $[-t_0,0]$, we first solve the equation without diffusion and then introduce the diffusion effects to our solutions.

When $D=0$, we can make the transformations

$$\begin{aligned}\widetilde{\sigma}_{12}(z,t) &= e^{-i\frac{g^2N}{c\Delta}z}e^{-i(k_0-k_c)z}\sigma_{12}(z,t),\\ \widetilde{E}(z,t) &= e^{-i\frac{g^2N}{c\Delta}z}E(z,t)\end{aligned} \qquad \text{(A2)}$$

and get the new equations

$$\begin{aligned}\partial_z\widetilde{E}(z,t) &= i\frac{g_{\text{eff}}N}{c}\widetilde{\sigma}_{12}(z,t),\\ \partial_t\widetilde{\sigma}_{12}(z,t) &= -i\eta z\widetilde{\sigma}_{12}(z,t) + ig_{\text{eff}}\widetilde{E}(z,t),\end{aligned} \qquad \text{(A3)}$$

where $g_{\text{eff}} = g\Omega_c/\Delta$. Following the method given in Ref. [9] and using the boundary conditions $\widetilde{\sigma}_{12}(z,t\to-\infty)=0$ and $\widetilde{E}(z=-L,t<0)=\widetilde{f}_{\text{in}}(t)$, we integrate the first equation and substitute it in the second one. Making use of Fourier transformation when the pulse is fully inside the medium, for t greater than the pulse length, we find

$$\widetilde{E}(k,t) = \widetilde{f}_{\text{in}}\left(\frac{k}{\eta}+\frac{\beta}{\eta L}+t\right)\left|\frac{k}{\eta}\right|^{-i\beta-1}G(\eta,\beta,L) \qquad \text{(A4)}$$

and

$$G(\eta,\beta,L) = \frac{1}{\eta}\beta e^{-\pi|\beta|/2}\sinh(\pi|\beta|)|\eta L|^{-i\beta}\Gamma(i\beta),$$

where $\widetilde{E}(k,t)=\int\widetilde{E}(z,t)e^{-ikz}dz$, $\beta=\frac{g_{\text{eff}}^2N}{\eta c}$ is the optical depth and we assume β is sufficiently large, $\Gamma(i\beta)$ is the Γ function, and $\widetilde{f}_{\text{in}}(t)=f_{\text{in}}(t)e^{i\frac{g^2N}{c\Delta}L}$ is the input pulse. According to the Maxwell-Bloch equations, we have $\widetilde{\sigma}_{12}(k,t)=\frac{k\cdot c}{g_{\text{eff}}N}\widetilde{E}(k,t)$.

We transform $\widetilde{\sigma}_{12}(k,t)$ back to $\sigma_{12}(k,t)$,

$$\begin{aligned}\sigma_{12}(k,t) &= f_{\text{in}}\left(\frac{k-k_i}{\eta}+t\right)e^{i\frac{g^2N}{c\Delta}L}\left|\frac{k-k_i}{\eta}-\frac{\beta}{\eta L}\right|^{-i\beta}\\ &\quad\times\text{sgn}\left(\frac{k-k_i}{\eta}-\frac{\beta}{\eta L}\right)\frac{c}{g_{\text{eff}}N}G \qquad \text{(A5)}\end{aligned}$$

with $k_i=\frac{g^2N}{c\Delta}+k_0-k_c-\frac{\beta}{L}$.

Now we introduce the diffusion; for the short time interval $[t,t+\Delta t]$, diffusion will cause a decay $e^{-Dk^2\Delta t}$ to $\sigma(k,t)$, or equally, there will be a decay $e^{-Dk^2\Delta t}$ on the signal $f_{\text{in}}(t')$ with $k=k_i-\eta(t-t')$. So the total decay during the write-in process for $f_{\text{in}}(t')$ is

$$e^{-D\int_{t'}^{0}[k_i-\eta(t-t')]^2dt} = e^{\frac{-D}{3\eta}[k_i^3-(k_i+\eta t')^3]}.$$

Thus, the solution for σ_{12} at $t=0$ is

$$\begin{aligned}\sigma_{12}(k,0) &= e^{\frac{-D}{3\eta}(k_i^3-k^3)}f_{\text{in}}\left(\frac{k-k_i}{\eta}\right)e^{i\frac{g^2N}{c\Delta}L}\left|\frac{k-k_i}{\eta}-\frac{\beta}{\eta L}\right|^{-i\beta}\\ &\quad\times\text{sgn}\left(\frac{k-k_i}{\eta}-\frac{\beta}{\eta L}\right)\frac{c}{g_{\text{eff}}N}G. \qquad \text{(A6)}\end{aligned}$$

We have assumed that the bandwidth of the memory is larger than the bandwidth of the input signal, $|\eta L|\gg\Delta\omega_s$, and the optical depth is sufficient large, $\beta\gtrsim 1$. The signal will be absorbed near $z=0$, and $\sigma_{12}(z,0)$ and $E(z,0)$ are nonzero only near $z=0$, so we can treat L as infinity during $[0,t_H]$. Also notice that the gradient is turned off during $[0,t_H]$, and the spatial frequency k will hold its value.

To get the solution in $[0,t_H]$, we solve Eq. (A1) with initial condition $\sigma_{12}(k,t=0)$ for σ_{12} and open boundary condition for E. In k space, we find

$$\sigma_{12}(k,t_H) = e^{-Dk^2t}e^{i\frac{g_{\text{eff}}^2N}{c}\frac{1}{k-\bar{k}}t_H}\sigma_{12}(k,0), \qquad \text{(A7)}$$

where $\bar{k}=\frac{g^2N}{c\Delta}+k_0-k_c$. Notice that $\sigma_{12}(k,t_H)$ gets a phase $e^{i\frac{g_{\text{eff}}^2N}{c}\frac{1}{k-\bar{k}}t_H}$, so the group velocity for $\sigma_{12}(z,t)$ is $v_g(k)=\frac{g_{\text{eff}}^2N}{c(k-\bar{k})^2}$. If the memory broadening $|\eta L|$ is not much larger than the signal pulse bandwidth, the spin wave $\sigma_{12}(z,t)$ will be nonzero near the ensemble boundary. Then the spin wave will propagate to the boundary and be reflected; this may ruin the spin wave coherence near the boundary and lower the memory efficiency. One way to avoid this effect is turning off the control field during storage, which makes the effective coupling $g_{\text{eff}}=0$ and the group velocity $v_g=0$.

To find the values for σ_{12} and E in the duration $[t_H,t_H+t_0]$, one needs to solve a modified version of Eq. (A1) where the sign of $i\eta z$ is reversed. We follow the method given in Ref. [9] and propagate these equation backward with final conditions $E(z=L,t>t_H)=f_{\text{out}}(t),\sigma_{12}(z,t\to\infty)=0$. Similar to the

write-in process, at time t_H, we have

$$\sigma_{12}(k,t_H) = e^{\frac{D}{3\eta}(k_i^3-k^3)} f_{\text{out}}\left(t_H + \frac{k-k_i}{-\eta}\right) \times e^{-i\frac{g^2N}{c\Delta}L}\left|\frac{k-k_i}{\eta} - \frac{\beta}{\eta L}\right|^{-i\beta} \times \text{sgn}\left(\frac{k-k_i}{\eta} - \frac{\beta}{\eta L}\right)\frac{c}{g_{\text{eff}}N}G^*. \quad \text{(A8)}$$

By matching the two solutions for σ_{12} at t_H Eqs. (A7) and (A8), we get

$$f_{\text{out}}(t_H + t) = d_W(t)d_H d_R(t) f_{\text{in}}(-t)\bar{G}, \quad \text{(A9)}$$

where

$$\bar{G} = \left|\eta L\left(t + \frac{\beta}{\eta L}\right)\right|^{-i2\beta} e^{i\frac{2Lg^2N}{c\Delta}} e^{-i\frac{g_{\text{eff}}^2 N}{c\eta(t+\frac{\beta}{\eta L})}t_H}\Gamma(i\beta)/\Gamma(-i\beta)$$

is a phase factor, $d_W(t) = e^{\frac{-D}{3\eta}[k_i^3-(k_i-\eta t)^3]}$, $d_H = e^{-D(k_i-\eta t)^2 t_H}$, and $d_R(t) = e^{\frac{-D}{3\eta}[k_i^3-(k_i-\eta t)^3]}$ are the diffusion decays for the write-in process $[-t_0,0]$, storage time $[0,t_H]$, and read-out process $[t_H,t_H+t_0]$ respectively.

Notice that the temporal shape is narrowed due to diffusion, but typically this change is very small, and the corresponding increase in pulse bandwidth is also small. Specifically, the increased pulse bandwidth is about $\frac{1}{\alpha_H} = 1 + D\eta^2 t_p^2 t_H$ times the initial pulse bandwidth, typically $D\eta^2 t_p^2 t_H \ll 1$. So our assumption that the the medium bandwidth is much larger than the pulse width is always valid, and we did not consider the case when this increased pulse bandwidth does not fit the medium bandwidth.

APPENDIX B

The Maxwell-Bloch equation for the three-dimensional model is

$$\dot{\sigma}_{12}(\mathbf{r},t) = i\frac{g\Omega_c}{\Delta}e^{i(k_0-k_c)z}E(\mathbf{r},t) - (i\eta z)\sigma_{12}(\mathbf{r},t) + D\nabla^2\sigma_{12}(\mathbf{r},t), \quad \text{(B1)}$$

$$\frac{\partial}{\partial z}E(\mathbf{r},t) = i\frac{gN\Omega_c}{c\Delta}e^{-i(k_0-k_c)z}\sigma_{12}(\mathbf{r},t) + i\frac{g^2N}{c\Delta}E(\mathbf{r},t).$$

To solve these equations, we first transform transverse coordinates x,y to Fourier space k_x,k_y,

$$\dot{\sigma}_{12}(k_x,k_y,z,t) = -(i\eta z + \gamma_k)\sigma_{12}(k_x,k_y,z,t) + i\frac{g\Omega_c}{\Delta}e^{i(k_0-k_c)z}E(k_x,k_y,z,t) + D\nabla_z^2\sigma_{12}(k_x,k_y,z,t), \quad \text{(B2)}$$

$$\frac{\partial}{\partial z}E(k_x,k_y,z,t) = i\frac{gN\Omega_c}{c\Delta}e^{-i(k_0-k_c)z}\sigma_{12}(k_x,k_y,z,t) + i\frac{g^2N}{c\Delta}E(k_x,k_y,z,t),$$

where $\gamma_k = D(k_x^2+k_y^2)$. Now we make the following transformation:

$$\bar{\sigma}_{12}(k_x,k_y,z,t) = e^{\gamma_k t}\sigma_{12}(k_x,k_y,z,t), \quad \bar{E}(k_x,k_y,z,t) = e^{\gamma_k t}E(k_x,k_y,z,t), \quad \text{(B3)}$$

and then we have

$$\dot{\bar{\sigma}}_{12}(k_x,k_y,z,t) = -(i\eta z)\bar{\sigma}_{12}(k_x,k_y,z,t) + i\frac{g\Omega_c}{\Delta}e^{i(k_0-k_c)z}\bar{E}(k_x,k_y,z,t) + D\nabla_z^2\bar{\sigma}_{12}(k_x,k_y,z,t), \quad \text{(B4)}$$

$$\frac{\partial}{\partial z}\bar{E}(k_x,k_y,z,t) = i\frac{gN\Omega_c}{c\Delta}e^{-i(k_0-k_c)z}\bar{\sigma}_{12}(k_x,k_y,z,t) + i\frac{g^2N}{c\Delta}\bar{E}(k_x,k_y,z,t).$$

These are actually quasi-1D equations, so we can solve these equations by the method we used before, and the output field is

$$\bar{f}_{\text{out}}(k_x,k_y,t_H+t) = d_W(t)d_H d_R(t)\bar{f}_{\text{in}}(k_x,k_y,-t)\bar{G}.$$

We transform back to $f_{\text{out}}(k_x,k_y,t_H+t)$ and get

$$f_{\text{out}}(k_x,k_y,t_H+t) = d_W(t)d_H d_R(t)d_\perp(t)f_{\text{in}}(k_x,k_y,-t)\bar{G}, \quad \text{(B5)}$$

where $d_\perp(t) = e^{-2\gamma_k t}e^{-\gamma_k t_H}$ is the transverse diffusion decay.

[1] L. M. Duan, M. Lukin, I. Cirac, and P. Zoller, Nature (London) **414**, 413 (2001).

[2] E. Knill, R. Laflamme, G. J. Milburn *et al.*, Nature (London) **409**, 46 (2001).

[3] M. Fleischhauer, A. Imamoglu, and J. P. Marangos, Rev. Mod. Phys. **77**, 633 (2005).

[4] M. Fleischhauer and M. D. Lukin, Phys. Rev. Lett. **84**, 5094 (2000).

[5] J. Nunn, I. A. Walmsley, M. G. Raymer, K. Surmacz, F. C. Waldermann, Z. Wang, and D. Jaksch, Phys. Rev. A **75**, 011401(R) (2007).

[6] B. Kraus, W. Tittel, N. Gisin, M. Nilsson, S. Kröll, and J. I. Cirac, Phys. Rev. A **73**, 020302(R) (2006).

[7] N. Sangouard, C. Simon, M. Afzelius, and N. Gisin, Phys. Rev. A **75**, 032327 (2007).

[8] M. Afzelius, C. Simon, H. de Riedmatten, and N. Gisin, Phys. Rev. A **79**, 052329 (2009).

[9] G. Hetet, J. J. Longdell, A. L. Alexander, P. K. Lam, and M. J. Sellars, arXiv:quant-ph/0612169v2.

[10] G. Hetet, J. J. Longdell, A. L. Alexander, P. K. Lam, and M. J. Sellars, Phys. Rev. Lett. **100**, 23601 (2008).

[11] G. Hétet, Ph.D. thesis, Australian National University, 2008, http://photonics.anu.edu.au/theses/.

[12] A. I. Lvovsky, B. C. Sanders, and W. Tittel, Nat. Photon. **3**, 706 (2009).

[13] N. B. Phillips, A. V. Gorshkov, and I. Novikova, Phys. Rev. A **78**, 023801 (2008).
[14] J. J. Longdell, E. Fraval, M. J. Sellars, and N. B. Manson, Phys. Rev. Lett. **95**, 063601 (2005).
[15] A. Amari, A. Walther, M. Sabooni, M. Huang, S. Kröll, M. Afzelius, I. Usmani, B. Lauritzen, N. Sangouard, H. De Riedmatten *et al.*, J. Lumin. **130**, 1579 (2010).
[16] C. Clausen, I. Usmani, F. Bussières, N. Sangouard, M. Afzelius, H. de Riedmatten, and N. Gisin, Nature (London) **469**, 508 (2011).
[17] E. Saglamyurek, N. Sinclair, J. Jin, J. A. Slater, D. Oblak, F. Bussières, M. George, R. Ricken, W. Sohler, and W. Tittel, Nature (London) **469**, 512 (2011).
[18] H. P. Specht, C. Nölleke, A. Reiserer, M. Uphoff, E. Figueroa, S. Ritter, and G. Rempe, Nature (London) **473**, 190 (2011).
[19] Y. Wang, J. Minář, G. Hétet, and V. Scarani, Phys. Rev. A **85**, 013823 (2012).
[20] C. Nölleke, A. Neuzner, A. Reiserer, C. Hahn, G. Rempe, and S. Ritter, Phys. Rev. Lett. **110**, 140403 (2013).
[21] M. Hosseini, B. M. Sparkes, G. Campbell, P. K. Lam, and B. C. Buchler, Nat. Commun. **2**, 174 (2011).
[22] D. B. Higginbottom, B. M. Sparkes, M. Rancic, O. Pinel, M. Hosseini, P. K. Lam, and B. C. Buchler, Phys. Rev. A **86**, 023801 (2012).
[23] O. Firstenberg, M. Shuker, R. Pugatch, D. R. Fredkin, N. Davidson, and A. Ron, Phys. Rev. A **77**, 043830 (2008).
[24] O. Firstenberg, M. Shuker, N. Davidson, and A. Ron, Phys. Rev. Lett. **102**, 43601 (2009).
[25] O. Firstenberg, P. London, D. Yankelev, R. Pugatch, M. Shuker, and N. Davidson, Phys. Rev. Lett. **105**, 183602 (2010).
[26] F. D. Walls and G. J. Milburn, *Quantum Optics* (Springer-Verlag, Berlin Heidelberg, 2008).
[27] T. M. Stace and A. N. Luiten, Phys. Rev. A **81**, 033848 (2010).
[28] G. R. Dennis, J. J. Hope, and M. T. Johnsson, Comput. Phys. Commun. **184**, 201 (2013).
[29] D. A. Steck, Rubidium 87 D Line Data, available online http://steck.us/alkalidata/ (Version 2.1.4, 2010).
[30] W. Happer, Rev. Mod. Phys. **44**, 169 (1972).